

BEST AVAILABLE COPY

**RAW SEQUENCE LISTING**

The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.

Application Serial Number: 10/591,271  
Source: JEP  
Date Processed by STIC: 9/18/06

***ENTERED***



IFWP

## RAW SEQUENCE LISTING

DATE: 09/18/2006

PATENT APPLICATION: US/10/591,271

TIME: 15:14:31

Input Set : A:\WIBL-PWO-035.TXT

Output Set: N:\CRF4\09182006\J591271.raw

4 <110> APPLICANT: Whitehead Institute for Biomedical Research  
 5 Harbison, Christopher  
 6 Young, Richard A.  
 7 Gordon, David B.  
 8 Fraenkel, Ernest  
 10 <120> TITLE OF INVENTION: BIOLOGICALLY-ACTIVE DNA-BINDING SITES  
 11 AND RELATED METHODS  
 13 <130> FILE REFERENCE: WIBL-PWO-035  
 C--> 15 <140> CURRENT APPLICATION NUMBER: US/10/591,271  
 C--> 16 <141> CURRENT FILING DATE: 2006-08-31  
 18 <150> PRIOR APPLICATION NUMBER: US 60/550,074  
 19 <151> PRIOR FILING DATE: 2004-03-04  
 21 <150> PRIOR APPLICATION NUMBER: US 60/604,470  
 22 <151> PRIOR FILING DATE: 2004-08-25  
 24 <160> NUMBER OF SEQ ID NOS: 59  
 26 <170> SOFTWARE: FastSEQ for Windows Version 4.0  
 28 <210> SEQ ID NO: 1  
 29 <211> LENGTH: 12  
 30 <212> TYPE: DNA  
 31 <213> ORGANISM: S. cerevisiae  
 33 <400> SEQUENCE: 1  
 34 atcgacgtg at 12  
 36 <210> SEQ ID NO: 2  
 37 <211> LENGTH: 12  
 38 <212> TYPE: DNA  
 39 <213> ORGANISM: S. paradoxus  
 41 <400> SEQUENCE: 2  
 42 atttcacatg at 12  
 44 <210> SEQ ID NO: 3  
 45 <211> LENGTH: 12  
 46 <212> TYPE: DNA  
 47 <213> ORGANISM: S. mikatae  
 49 <400> SEQUENCE: 3  
 50 atactacgtg ac 12  
 52 <210> SEQ ID NO: 4  
 53 <211> LENGTH: 12  
 54 <212> TYPE: DNA  
 55 <213> ORGANISM: S. bayanus  
 57 <400> SEQUENCE: 4  
 58 cttgcacgtg cc 12  
 60 <210> SEQ ID NO: 5  
 61 <211> LENGTH: 10  
 62 <212> TYPE: DNA

(pg. 6)

## RAW SEQUENCE LISTING

DATE: 09/18/2006

PATENT APPLICATION: US/10/591,271

TIME: 15:14:31

Input Set : A:\WIBL-PWO-035.TXT

Output Set: N:\CRF4\09182006\J591271.raw

```

63 <213> ORGANISM: Artificial Sequence
65 <220> FEATURE:
66 <223> OTHER INFORMATION: Recognition sequence for YDR026C
68 <400> SEQUENCE: 5
69 ttaccgcm                               10
71 <210> SEQ ID NO: 6
72 <211> LENGTH: 18
73 <212> TYPE: DNA
74 <213> ORGANISM: Artificial Sequence
76 <220> FEATURE:
77 <223> OTHER INFORMATION: Unlabeled competitor for Cin5 binding
79 <400> SEQUENCE: 6
80 ggcacattac ctaagggc                     18
82 <210> SEQ ID NO: 7
83 <211> LENGTH: 18
84 <212> TYPE: DNA
85 <213> ORGANISM: Artificial Sequence
87 <220> FEATURE:
88 <223> OTHER INFORMATION: Unlabeled competitor for Cin5 binding
90 <400> SEQUENCE: 7
91 ggcacattac taaagggc                     18
93 <210> SEQ ID NO: 8
94 <211> LENGTH: 13
95 <212> TYPE: DNA
96 <213> ORGANISM: Artificial Sequence
98 <220> FEATURE:
99 <223> OTHER INFORMATION: Discovered binding site for Abf1
W--> 101 <221> NAME/KEY: misc_feature
102 <222> LOCATION: 7, 8, 9, 10
103 <223> OTHER INFORMATION: n = A,T,C or G
W--> 105 <400> 8
W--> 106 rtcaaytnnnn acg                     13
108 <210> SEQ ID NO: 9
109 <211> LENGTH: 14
110 <212> TYPE: DNA
111 <213> ORGANISM: Artificial Sequence
113 <220> FEATURE:
114 <223> OTHER INFORMATION: Known binding site for Abf1
W--> 116 <221> NAME/KEY: misc_feature
117 <222> LOCATION: 7, 8, 9, 10
118 <223> OTHER INFORMATION: n = A,T,C or G
W--> 120 <400> 9
W--> 121 rtcaaytnnnn acgw                     14
123 <210> SEQ ID NO: 10
124 <211> LENGTH: 14
125 <212> TYPE: DNA
126 <213> ORGANISM: Artificial Sequence
128 <220> FEATURE:
129 <223> OTHER INFORMATION: Known binding site for Abf2

```

## RAW SEQUENCE LISTING

DATE: 09/18/2006

PATENT APPLICATION: US/10/591,271

TIME: 15:14:31

Input Set : A:\WIBL-PWO-035.TXT

Output Set: N:\CRF4\09182006\J591271.raw

```

131 <400> SEQUENCE: 10
132 aaagtgcacc catt                                     14
134 <210> SEQ ID NO: 11
135 <211> LENGTH: 18
136 <212> TYPE: DNA
137 <213> ORGANISM: Artificial Sequence
139 <220> FEATURE:
140 <223> OTHER INFORMATION: Discovered binding site for Azf1
142 <400> SEQUENCE: 11
143 ywttkckkty yckgykky                                 18
145 <210> SEQ ID NO: 12
146 <211> LENGTH: 10
147 <212> TYPE: DNA
148 <213> ORGANISM: Artificial Sequence
150 <220> FEATURE:
151 <223> OTHER INFORMATION: Discovered binding site for Cad1
153 <400> SEQUENCE: 12
154 mttastmakc                                           10
156 <210> SEQ ID NO: 13
157 <211> LENGTH: 19
158 <212> TYPE: DNA
159 <213> ORGANISM: Artificial Sequence
161 <220> FEATURE:
162 <223> OTHER INFORMATION: Known binding site for Dal81
164 <400> SEQUENCE: 13
165 aaaagccgcg ggcgggatt                                 19
167 <210> SEQ ID NO: 14
168 <211> LENGTH: 17
169 <212> TYPE: DNA
170 <213> ORGANISM: Artificial Sequence
172 <220> FEATURE:
173 <223> OTHER INFORMATION: Discovered binding site for Gal4
W--> 175 <221> NAME/KEY: misc_feature
176 <222> LOCATION: 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14
177 <223> OTHER INFORMATION: n = A,T,C or G
W--> 179 <400> 14
W--> 180 cggnnnnnnnn nnnnccg                             17
182 <210> SEQ ID NO: 15
183 <211> LENGTH: 17
184 <212> TYPE: DNA
185 <213> ORGANISM: Artificial Sequence
187 <220> FEATURE:
188 <223> OTHER INFORMATION: Known binding site for Gal4
W--> 190 <221> NAME/KEY: misc_feature
191 <222> LOCATION: 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14
192 <223> OTHER INFORMATION: n = A,T,C or G
W--> 194 <400> 15
W--> 195 cggnnnnnnnn nnnnccg                             17
197 <210> SEQ ID NO: 16

```

## RAW SEQUENCE LISTING

DATE: 09/18/2006

PATENT APPLICATION: US/10/591,271

TIME: 15:14:31

Input Set : A:\WIBL-PWO-035.TXT

Output Set: N:\CRF4\09182006\J591271.raw

```

198 <211> LENGTH: 17
199 <212> TYPE: DNA
200 <213> ORGANISM: Artificial Sequence
202 <220> FEATURE:
203 <223> OTHER INFORMATION: Known binding site for Gal80
W--> 205 <221> NAME/KEY: misc_feature
206 <222> LOCATION: 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14
207 <223> OTHER INFORMATION: n = A,T,C or G
W--> 209 <400> 16
W--> 210 cggnnnnnnnn nnnnccg 17
212 <210> SEQ ID NO: 17
213 <211> LENGTH: 11
214 <212> TYPE: DNA
215 <213> ORGANISM: Artificial Sequence
217 <220> FEATURE:
218 <223> OTHER INFORMATION: Known binding site for Gln3
220 <400> SEQUENCE: 17
221 gataagataa g 11
223 <210> SEQ ID NO: 18
224 <211> LENGTH: 11
225 <212> TYPE: DNA
226 <213> ORGANISM: Artificial Sequence
228 <220> FEATURE:
229 <223> OTHER INFORMATION: Discovered binding site for Hap1
W--> 231 <221> NAME/KEY: misc_feature
232 <222> LOCATION: 3, 4, 8
233 <223> OTHER INFORMATION: n = A,T,C or G
W--> 235 <400> 18
W--> 236 gggnnatancg s 11
238 <210> SEQ ID NO: 19
239 <211> LENGTH: 13
240 <212> TYPE: DNA
241 <213> ORGANISM: Artificial Sequence
243 <220> FEATURE:
244 <223> OTHER INFORMATION: Known binding site for Hap4
W--> 246 <221> NAME/KEY: misc_feature
247 <222> LOCATION: 3, 4, 10, 12
248 <223> OTHER INFORMATION: n = A,T,C or G
W--> 250 <400> 19
W--> 251 ycnmccaatn anm 13
253 <210> SEQ ID NO: 20
254 <211> LENGTH: 12
255 <212> TYPE: DNA
256 <213> ORGANISM: Artificial Sequence
258 <220> FEATURE:
259 <223> OTHER INFORMATION: Known binding site for HaeI
261 <400> SEQUENCE: 20
262 kgmcagcgtg tc 12
264 <210> SEQ ID NO: 21

```

## RAW SEQUENCE LISTING

DATE: 09/18/2006

PATENT APPLICATION: US/10/591,271

TIME: 15:14:31

Input Set : A:\WIBL-PWO-035.TXT

Output Set: N:\CRF4\09182006\J591271.raw

```

265 <211> LENGTH: 12
266 <212> TYPE: DNA
267 <213> ORGANISM: Artificial Sequence
269 <220> FEATURE:
270 <223> OTHER INFORMATION: Known binding site for Hap1
W--> 272 <221> NAME/KEY: misc_feature
273 <222> LOCATION: 4, 5, 6, 9
274 <223> OTHER INFORMATION: n = A,T,C or G
W--> 276 <400> 21
W--> 277 cggnntanc gg 12
279 <210> SEQ ID NO: 22
280 <211> LENGTH: 13
281 <212> TYPE: DNA
282 <213> ORGANISM: Artificial Sequence
284 <220> FEATURE:
285 <223> OTHER INFORMATION: Discovered binding site for Hsfl
W--> 287 <221> NAME/KEY: misc_feature
288 <222> LOCATION: 5, 6, 7, 8, 9, 10
289 <223> OTHER INFORMATION: n = A,T,C or G
W--> 291 <400> 22
W--> 292 ttcynnnnnn ttc 13
294 <210> SEQ ID NO: 23
295 <211> LENGTH: 14
296 <212> TYPE: DNA
297 <213> ORGANISM: Artificial Sequence
299 <220> FEATURE:
300 <223> OTHER INFORMATION: Known binding site for Hsfl
W--> 302 <221> NAME/KEY: misc_feature
303 <222> LOCATION: 9, 10
304 <223> OTHER INFORMATION: n = A,T,C or G
W--> 306 <400> 23
W--> 307 ttctagaann ttct 14
309 <210> SEQ ID NO: 24
310 <211> LENGTH: 11
311 <212> TYPE: DNA
312 <213> ORGANISM: Artificial Sequence
314 <220> FEATURE:
315 <223> OTHER INFORMATION: Discovered binding site for Imel
W--> 317 <221> NAME/KEY: misc_feature
318 <222> LOCATION: 8
319 <223> OTHER INFORMATION: n = A,T,C or G
W--> 321 <400> 24
W--> 322 aakgaaankw a 11
324 <210> SEQ ID NO: 25
325 <211> LENGTH: 10
326 <212> TYPE: DNA
327 <213> ORGANISM: Artificial Sequence
329 <220> FEATURE:
330 <223> OTHER INFORMATION: Known binding site for Ino2

```

## RAW SEQUENCE LISTING ERROR SUMMARY

DATE: 09/18/2006

PATENT APPLICATION: US/10/591,271

TIME: 15:14:32

Input Set : A:\WIBL-PWO-035.TXT

Output Set: N:\CRF4\09182006\J591271.raw

**Please Note:**

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:8; N Pos. 7,8,9,10

Seq#:9; N Pos. 7,8,9,10

Seq#:14; N Pos. 4,5,6,7,8,9,10,11,12,13,14

Seq#:15; N Pos. 4,5,6,7,8,9,10,11,12,13,14

Seq#:16; N Pos. 4,5,6,7,8,9,10,11,12,13,14

Seq#:18; N Pos. 3,4,8

Seq#:19; N Pos. 3,4,10,12

Seq#:21; N Pos. 4,5,6,9

Seq#:22; N Pos. 5,6,7,8,9,10

Seq#:23; N Pos. 9,10

Seq#:24; N Pos. 8

Seq#:29; N Pos. 3,7,8

Seq#:30; N Pos. 10,11

Seq#:32; N Pos. 3,7,8

Seq#:35; N Pos. 4,5,6,7,8,9,10,11,12,13

Seq#:39; N Pos. 7,8

Seq#:44; N Pos. 4,6

Seq#:50; N Pos. 4

Seq#:51; N Pos. 6,7,8

Seq#:54; N Pos. 7,8

Seq#:55; N Pos. 7

Seq#:56; N Pos. 4,5,6,7

## VERIFICATION SUMMARY

DATE: 09/18/2006

PATENT APPLICATION: US/10/591,271

TIME: 15:14:32

Input Set : A:\WIBL-PWO-035.TXT

Output Set: N:\CRF4\09182006\J591271.raw

L:15 M:270 C: Current Application Number differs, Replaced Current Application Number  
L:16 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:101 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:105 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:8  
L:106 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:0  
L:116 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:120 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:9  
L:121 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:0  
L:175 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:179 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:14  
L:180 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:0  
L:190 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:194 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:15  
L:195 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:0  
L:205 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:209 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:16  
L:210 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:0  
L:231 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:235 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:18  
L:236 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18 after pos.:0  
L:246 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:250 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:19  
L:251 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19 after pos.:0  
L:272 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:276 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:21  
L:277 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21 after pos.:0  
L:287 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:291 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:22  
L:292 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22 after pos.:0  
L:302 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:306 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:23  
L:307 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23 after pos.:0  
L:317 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:321 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:24  
L:322 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:24 after pos.:0  
L:376 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:380 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:29  
L:381 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29 after pos.:0  
L:391 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:395 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:30  
L:396 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30 after pos.:0  
L:417 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:421 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:32  
L:422 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:32 after pos.:0  
L:454 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:458 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:35  
L:459 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:35 after pos.:0  
L:502 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!



**VERIFICATION SUMMARY**

DATE: 09/18/2006

PATENT APPLICATION: US/10/591,271

TIME: 15:14:32

Input Set : A:\WIBL-PWO-035.TXT

Output Set: N:\CRF4\09182006\J591271.raw

L:506 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:39  
L:507 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:39 after pos.:0  
L:561 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:565 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:44  
L:566 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:44 after pos.:0  
L:631 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:635 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:50  
L:636 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:50 after pos.:0  
L:646 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:650 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:51  
L:651 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:51 after pos.:0  
L:683 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:687 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:54  
L:688 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:54 after pos.:0  
L:698 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:702 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:55  
L:703 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:55 after pos.:0  
L:713 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:717 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:56  
L:718 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:56 after pos.:0

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**